Sample Final Exam B – Part A

1. A consumer group surveyed the prices for white cotton twin sheet sets in five different department stores and reported the median price to be \$37. We visited four of the five stores, and found the prices to be \$37, \$42, \$37 and \$35. Assuming that the consumer group is correct, what is the price of the sheet set at the store we did not visit?

   (A) \$34

   (B) \$35

   (C) \$36

   (D) \$37

   (E) We dont know – the price at the store we didnt visit could be anything.

2. The five-number summary of the heights (in inches) of the players on a basketball team is calculated to be:

72 76 78 79 81

The distribution of heights for the team is:

(A) skewed to the left and so the mean is greater than the median.
(B) skewed to the right and so the mean is greater than the median.
(C) skewed to the left and so the median is greater than the mean.
(D) skewed to the right and so the median is greater than the mean.
(E) approximately symmetric and so the mean and median are approximately equal.

3. The University of Manitoba uses a grade point system with a maximum possible Grade Point Average of 4.5. For each course, a student is awarded the following number of grade points per credit hour for each of the possible letter grades:

Letter Grade	Grade Points
A^+	4.5
A	4.0
B^+	3.5
B	3.0
C^+	2.5
C	2.0
D	1.0
F	0.0

One year, a student took the following courses, and received the letter grades shown:

Course	Credit Hours	Letter Grade
CIVL 2830	2	B^+
HIST 1260	3	A^+
ECE 2130	4	A
ECE 2160	5	C
COMP 2061	6	B

What is the student's GPA for the year?

(A) 3.265 (B) 3.375 (C) 3.400 (D) 3.225 (E) 3.175

4. Over a period of one month, the correlation between the daily high temperature X in Winnipeg and the daily high temperature Y in Grand Forks, North Dakota (both measured in degrees Celsius) is calculated to be 0.79. If we had measured the temperatures in degrees Fahrenheit (temperature in $^\circ$F $= 32 + 1.8$ (temperature in $^\circ$C)), the correlation between X and Y would be equal to:

(A) 0.79 (B) 0.44 (C) 0.89 (D) 0.68 (E) 0.72

5. We would like to examine how the age (X) of a certain model of car affects its selling price (Y). The age (in years) and price (in \$) for a sample of 15 cars of the same make and model are recorded from the classified ads in the newspaper one weekend. A scatterplot shows a linear relationship between X and Y and the correlation is calculated to be $r = -0.9$. Despite the strong correlation between age and price, we cannot say that a car getting older **causes** its value to decrease. This is because of the potential effect of one or more lurking variables. Which of the following is a likely lurking variable in this case?

   (A) colour

   (B) make and model

   (C) mileage

   (D) age

   (E) all of the above

6. An experiment will be conducted to determine how the length of an icicle is affected by temperature and wind speed. Two different temperatures (−20°C and −30°C) and three different wind speeds (20 km/h, 40km/h and 60km/h) will be examined. Ten icicles will be subjected to each combination of factor levels.

   What are the treatments in this experiment?

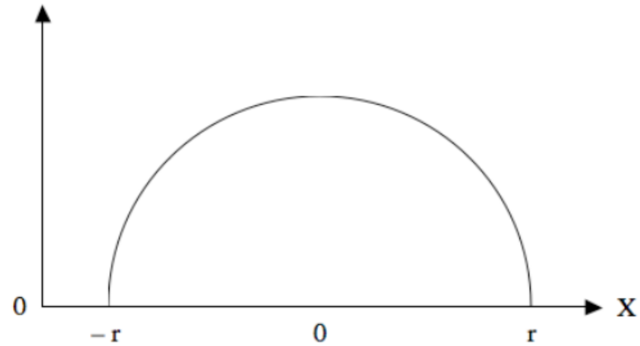
   (A) temperature and wind speed
   (B) temperature, wind speed and icicle length
   (C) −20°C, −30°, 20 km/h, 40km/h, 60km/h
   (D) −20°C/20 km/h, −30°/20 km/h, −20°C/40 km/h, −30°/40 km/h, −20°C/60 km/h, −30°/60 km/h
   (E) temperature, wind speed and number of icicles

7. A sample of 100 University of Manitoba professors is asked whether their spouse also works in the education field. What is the population of interest?

   (A) the 100 University of Manitoba professors
   (B) all University of Manitoba professors
   (C) the spouses of the 100 University of Manitoba professors
   (D) the spouses of all University of Manitoba professors
   (E) all University of Manitoba professors whose spouses also work in the education field

The next **two** questions (**8** to **9**) refer to the following:

A random variable X is described by a semi-circular density curve with mean 0 and standard deviation 0.4, as shown below:



8. What must be the value of r in order for this to be a legitimate density curve? (Hint: Recall the area of the circle is equal to πr^2, where r is the radius).

(A) $\sqrt{\frac{1}{\pi}}$ (B) $\sqrt{2\pi}$ (C) $\sqrt{\frac{\pi}{2}}$ (D) $\sqrt{\pi}$ (E) $\sqrt{\frac{2}{\pi}}$

9. We take a simple random sample of 100 individuals from the semi-circular distribution and calculate the sample mean $\bar{x}$. The sampling distribution of $\bar{X}$ is:

(A) approximately normal with mean 0 and standard deviation 0.004.

(B) semi-circular with mean 0 and standard deviation 0.04.

(C) approximately normal with mean 0 and standard deviation 0.4.

(D) semi-circular with mean 0 and standard deviation 0.004.

(E) approximately normal with mean 0 and standard deviation 0.04.

10. The scores on a university examination are normally distributed with a mean of 62 and a standard deviation of 11. If the bottom 5% of the students get failing grades on the examination, what is the lowest mark that a student can obtain and still be awarded a passing grade?

(A) 51 (B) 57 (C) 44 (D) 40 (E) 50

11. The number of strokes X taken by professional golfers on a particular hole follows a probability distribution as shown below, where k is some constant:

x	3	4	5	6	7
$P(X = x)$	0.12	$4k$	$3k$	$2k$	0.07

What proportion of golfers take at least five strokes on this hole?

(A) 0.48 (B) 0.25 (C) 0.57 (D) 0.43 (E) 0.52

12. A student driving to university must pass through seven sets of traffic lights. Suppose it is known that each of the traffic lights is red 35% of the time and that all lights function independently. What is the probability that the student will have to stop at two or more lights on her way to university?

(A) 0.6828 (B) 0.2985 (C) 0.4893 (D) 0.7662 (E) 0.5997

The next **two** questions (**13** to **14**) refer to the following:

The fill volumes per bottle for a certain brand of beer follow a normal distribution with mean 341 ml and standard deviation 3 ml.

13. If you buy a case of 24 bottles of beer, what is the probability that the bottles will contain an average between 340 and 342 ml? (Assume that the 24 bottles can be considered a simple random sample.)

(A) 0.7359 (B) 0.8968 (C) 0.6723 (D) 0.9252 (E) 0.5870

14. If you buy a twelve-pack of beer, what is the probability that the total volume of beer will exceed 4.1 liters? (Assume that the 12 bottles can be considered a simple random sample.)

(A) 0.0838 (B) 0.2206 (C) 0.1397 (D) 0.4443 (E) 0.3632

15. We would like to estimate the true mean contents μ of a bottle of a certain brand of cough syrup. We measure the contents of a random sample of 35 bottles and we calculate a mean of 252 ml. It is known that the contents of a bottle of cough syrup follow a normal distribution with standard deviation 1.5 ml. An 82% confidence interval for μ is:

(A) $2.52 \pm 1.48 \left(\frac{1.5}{\sqrt{35}} \right)$

(B) $2.52 \pm 1.34 \left(\frac{1.5}{\sqrt{35}} \right)$

(C) $2.52 \pm 0.82 \left(\frac{1.5}{\sqrt{35}} \right)$

(D) $2.52 \pm 0.92 \left(\frac{1.5}{\sqrt{35}} \right)$

(E) $2.52 \pm 1.28 \left(\frac{1.5}{\sqrt{35}} \right)$

16. A random variable X follows a normal distribution with known standard deviation σ. We would like to construct a confidence interval for the true mean μ of the distribution of X. For which of the following combinations of sample size and confidence level would the confidence interval be the widest?

(A) 97% confidence level with $n = 25$

(B) 97% confidence level with $n = 100$

(C) 99% confidence level with $n = 25$

(D) 99% confidence level with $n = 100$

(E) depends on the value of σ.

17. A random variable X follows a normal distribution with standard deviation 5. We take a random sample of 100 individuals from the population and calculate a confidence interval for μ to be (40.973, 43.027). What is the confidence level for this interval?

(A) 90% (B) 95% (C) 96% (D) 98% (E) 99%

The next **two** questions (**18** to **19**) refer to the following:

City engineers in Hamilton, Ontario would like to estimate the true mean commuting distance of all workers in the city between home and their principal place of business. They calculate that, in order to estimate this mean to within ± 1 kilometer with 99% confidence, they require a sample of 120 people.

18. What sample size would be required to estimate the true mean commuting distance for all workers in Hamilton to within ± 2 kilometers with 99% confidence?

(A) 30 (B) 60 (C) 85 (D) 240 (E) 480

19. The city of Montreal has a population five times greater than that of Hamilton (and so it would be logical to assume that it has five times as many workers). Suppose we wanted to estimate the true mean commuting distance between home and work for all workers in Montreal to within ± 1 kilometer with 99% confidence. Assuming equal standard deviations for the two cities, we would require a sample of how many Montreal workers?

(A) 24 (B) 120 (C) 269 (D) 600 (E) 3000

20. The level of significance α of a hypothesis test is:

(A) the lowest P-value for which the null hypothesis will be rejected.

(B) the probability that the null hypothesis will be rejected.

(C) the lowest value of the test statistic for which the null hypothesis will be rejected.

(D) the probability that the null hypothesis is true.

(E) the highest P-value for which the null hypothesis will be rejected.

21. A man accused of committing a crime is taking a polygraph (lie detector) test. The polygraph is essentially testing the hypotheses

$$H_0\text{: The man is telling the truth. vs. } H_a\text{: The man is lying.}$$

Suppose we use a 5% level of significance. Based on the man's responses to the questions asked, the polygraph determines a P-value of 0.08. We conclude that:

(A) There is insufficient evidence that the man is telling the truth.

(B) There is sufficient evidence that the man is telling the truth.

(C) There is insufficient evidence that the man is lying.

(D) The probability that the man is lying is 0.33.

(E) The probability that the man is telling the truth is 0.33.

22. A clinic measures the blood calcium of a sample of 100 healthy pregnant young women. The mean and standard deviation of these 100 measurements are, respectively, 9.8 and 0.5. Calcium levels for healthy pregnant young women are known to follow a normal distribution with standard deviation 0.4. We would like to conduct a test of significance to determine whether the data provide evidence that the mean calcium level in the population of healthy pregnant young women is less than 10. The test statistic for the appropriate test is:

(A) $t = -5.00$ (B) $z = -4.00$ (C) $z = -5.00$ (D) $t = -4.00$ (E) $z = -0.40$

23. We would like to test whether the true mean amount of money spent by gamblers at a casino differs from \$100; that is, we want to test $H_0 : \mu = 100$ vs. $H_a : \mu \neq 100$. We select a random sample of 50 gamblers and calculate a 98% confidence interval to be (\$99, \$114). Which of the following is true?

(A) We would reject H_0 at the $\alpha = 0.02$ level of significance.

(B) We would not reject H_0 at the $\alpha = 0.04$ level of significance.

(C) We would not reject H_0 at the $\alpha = 0.02$ level of significance.

(D) We would reject H_0 at the $\alpha = 0.01$ level of significance.

(E) A confidence interval cannot be used to conduct this test.

24. Five runners were asked to run a 10-kilometre race on each of two consecutive weeks. In one of the races, the runners wear one brand of shoe and in the other a different brand (with order randomly determined). All runners are timed and are asked to run their best in each race. The results (in minutes) are given below, with some sample calculations that may be useful:

Runner	1	2	3	4	5	Mean	Std. Dev.
Brand 1	31.2	29.3	30.5	32.2	33.1	31.26	1.47
Brand 2	32.0	29.0	30.9	32.7	33.0	31.52	1.62
Difference	−0.8	0.3	−0.4	−0.5	0.1	−0.26	0.45

We wish to conduct a hypothesis test to determine if there is evidence that running times for the two brands differ. Suppose it is known that differences in times for the two brands follow a normal distribution. The P-value for the appropriate test of significance is:

(A) between 0.40 and 0.50

(B) between 0.20 and 0.30

(C) between 0.10 and 0.20

(D) between 0.05 and 0.10

(E) between 0.04 and 0.05

25. A manufacturer of electronic components will send a large shipment to a retailer only if there is significant evidence that less than 5% of the components in the shipment are defective. The manufacturer tests a random sample of 300 components and finds that 9 of them are defective. The test statistic for the appropriate test of $H_0 : p = 0.05$ vs. $H_a : p < 0.05$ is:

(A) $z = \dfrac{0.03 - 0.05}{\sqrt{\dfrac{(0.03)(0.97)}{300}}}$

(B) $z = \dfrac{0.05 - 0.03}{\sqrt{\dfrac{(0.05)(0.95)}{300}}}$

(C) $z = \dfrac{0.03 - 0.05}{\sqrt{\dfrac{(0.03)(0.05)}{300}}}$

(D) $z = \dfrac{0.03 - 0.05}{\sqrt{\dfrac{(0.05)(0.95)}{300}}}$

(E) $z = \dfrac{0.05 - 0.03}{\sqrt{\dfrac{(0.03)(0.97)}{300}}}$

26. You want to estimate the proportion of Canadians who would support a merger of the federal NDP and Liberal parties to within ± 0.035 with 92% confidence. How large a sample is required?

(A) 625 (B) 784 (C) 466 (D) 816 (E) 553

27. It is commonly believed that about 10% of the population is left-handed. One researcher believes that the actual proportion is lower than this. He takes a simple random sample of 200 individuals and finds that 16 of them (8%) are left handed. What are the hypotheses for the appropriate test of significance to test the researcher's suspicion?

(A) H_0: $\hat{p} = 0.08$ vs. H_a: $\hat{p} < 0.08$

(B) H_0: $\hat{p} = 0.10$ vs. H_a: $\hat{p} < 0.10$

(C) H_0: $p = \hat{p}$ vs. H_a: $p < \hat{p}$

(D) H_0: $p = 0.10$ vs. H_a: $p < 0.10$

(E) H_0: $p = 0.08$ vs. H_a: $p < 0.08$

Sample Final Exam B – Part B

1. Most experts agree that depression is best treated using a combination of medication and therapy. A psychiatrist plans to conduct an experiment to examine the effects of two different types of therapy (cognitive behavioural therapy and psychodynamic therapy) and three different medications (Zoloft, Paxil and Effexor). Effects of therapy and medication are believed to differ depending on the severity of a subject's depression so the psychiatrist decides to use a randomized block design. 180 subjects with mild depression and 120 subjects with severe depression are available for the experiment. At the end of one year, the improvement of each subject's condition will be evaluated.

(a) Identify each of the following in this experiment:

(i) Factor(s):

(ii) Treatment(s):

(iii) Response Variable:

(iv) Blocking Variable:

(b) How is randomization used in this experiment?

(c) If individuals receiving one of the treatments improve significantly more than others, can we conclude that the treatment may be the **cause**? Why or why not?

2. Event A has probability 0.4 to occur and Event B has probability 0.7 to occur.
   (a) Are A and B mutually exclusive (disjoint)? Explain.

   (b) If A and B are independent, what is the probability P(A or B^c)?

3. One measure of the quality of education provided by a university is the number of students per class. The president of a large university would like to estimate the true mean size of all third-year classes at the university. A random sample of 30 third-year classes is selected. The number of students in each of the classes are ordered and shown below:

10	11	14	14	15	17	18	20	22	22
22	23	26	26	27	28	31	33	34	34
36	42	44	49	50	55	62	68	77	82

From these data, the sample mean is calculated to be 33.73. The population standard deviation of class sizes is known to be 18.50.

(a) Before even looking at the data, we know that the distribution of class sizes could not possibly be normal. Explain why.

(b) Construct a stemplot for these data. What is the shape of the data distribution?

(c) It is clear that class size does not follow a normal distribution. Explain why it is still appropriate to use inference methods which rely on the assumption of normality.

(d) Conduct an appropriate hypothesis test, at the 1% level of significance, to determine whether there is evidence that the true mean third-year class size at this university is lower than the national average of 40. Show all of the steps, including the statement of hypotheses, the calculation of the appropriate test statistic and P-value, and a carefully-worded conclusion.

4. Sixteen people volunteered to be part of an experiment. All 16 people were Caucasian, between the ages of 25 and 35, and were supplied with nice clothes. Eight of the people were male and eight were female. The question of interest in this experiment was whether females receive faster service at restaurants than males. Each of the eight male participants was randomly assigned a restaurant, and each of the eight females was randomly assigned to one of these same eight restaurants. One Friday night, all 16 people went out to eat, each one alone. The male and female assigned to the same restaurant would arrive within five minutes of each other, with the order determined by flipping a coin (male first or female first). Each person then ordered a similar drink and a similar meal. The time (in minutes) until the food arrived at the table was recorded. Some information that may be helpful is shown below:

Females	Males	Difference (d = F−M)
mean = 17.5	mean = 20.2	mean =−2.7
std. dev. = 3.7	std. dev. = 4.8	std. dev. = 2.9

(a) Conduct an appropriate hypothesis test, at the 1% level of significance. Show all of the steps, including the statement of hypotheses, the calculation of the appropriate test statistic and P-value, and a carefully-worded conclusion. Assume the appropriate normality conditions are satisfied.

(b) Interpret the P-value of the test to someone with little or no background in statistics.